Interference in Joint Picture Naming

Citation for published version:
Gambi, C, Van de Cavey, J & Pickering, MJ 2015, 'Interference in Joint Picture Naming', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 41, no. 1.
https://doi.org/10.1037/a0037438

Digital Object Identifier (DOI):
10.1037/a0037438

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Journal of Experimental Psychology: Learning, Memory, and Cognition

Interference in Joint Picture Naming

Chiara Gambi

University of Edinburgh and Saarland University

Joris Van de Cavey

Universiteit Gent

Martin J. Pickering

University of Edinburgh

Address for correspondence:

Chiara Gambi

Psycholinguistics Group

Department of Computational Linguistics and Phonetics

Building C 7.4

Campus Saarbruecken

Saarland University

Saarbruecken, 66123, Germany

gambi@coli.uni-saarland.de

Phone: +49 (0)681 302 6554

Fax: +49 (0)681 302 6561

Abstract

In four experiments we showed that picture naming latencies are affected by beliefs about the task concurrently performed by another speaker. Participants took longer to name pictures when they believed that their partner concurrently named pictures than when they believed their partner was silent (Experiments 1 and 4) or concurrently categorized the pictures as being from the same or from different semantic categories (Experiment 2). However, picture naming latencies were not affected by beliefs about *what* one's partner said, as it did not matter whether participants believed their partner produced the same utterance, or an utterance that differed by ordering (Experiments 1 and 2) or lexical content (Experiments 3 and 4). These findings are consistent with the proposal that speakers represent whether another speaker is preparing to speak, but not what they are preparing to say.

Keywords: joint task; co-representation; actor-conflict; language production; picture naming

This is the authors' manuscript of the final accepted version of this article:

http://psycnet.apa.org/psycinfo/2014-31794-001/

Published in Journal of Experimental Psychology: Learning, Memory, & Cognition

Interference in Joint Picture Naming

Is the way in which people speak affected by their beliefs about other people's speech? If so, it would suggest that speakers' representations of their partners' speech are in the same format as speakers' representations of their own speech (as the former can affect the latter). It is often assumed that language production and comprehension share representations (the parity hypothesis; Pickering & Garrod, 2004), and that this explains effects of comprehension on production. However, we do not know whether people form representations of others' utterances when they are not comprehended but only imagined, and whether such representations are in the same format as representations of their own utterances. To address this question, we asked whether performance in a joint picture-naming task differs depending on whether the speaker is told that his partner is concurrently performing the same or a related task, or no task.

Evidence that comprehended utterances affect concurrent production comes, for example, from the picture-word interference paradigm. In this paradigm, participants name pictures while ignoring written (or auditory) distractor words. Responses are fastest when the distractor word is the picture's name. But more importantly, they are slower when the distractor is another word than when the distractor is a row of Xs (e.g., Glaser & Düngelhoff, 1984). This shows that comprehension of the distractor word affects production of the target word. But even if representations used during language comprehension are in the same format as representations used during language production, we do not know whether comprehending another person's utterance and imagining such an utterance have comparable effects on production.

If we treat language production as a form of action and language comprehension as a form of perception, then parity in language constitutes an example of common coding (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1997); i.e., the hypothesis that action and perception share a representational code. In addition, if we specifically treat language comprehension as a form of action perception (i.e., perception of others' actions), this hypothesis leads to the claim that representations of others' actions are in the same format as representations underlying one's own actions (Heyes, 2011; Knoblich, Butterfill, & Sebanz, 2011; Sebanz, Bekkering, & Knoblich, 2006; Wilson & Knoblich, 2005).

This predicts that representing the actions of others affects the planning and execution of one's own actions. This prediction has been confirmed by several studies. First, studies using individual tasks (i.e., tasks in which one participant is acting alone) showed that perceiving an action activates the motor representations underlying that action, in a way that is automatic and interferes with execution of a different action (see Heyes, 2011 for a review). This is analogous to findings from picture-word interference studies. Second, studies using joint tasks showed how participants "acting together" automatically represent their partner's task as well as their own (see Knoblich, et al., 2011 for a review), even when it is not relevant for their own task, and when doing so interferes with the execution of their own task.

Importantly for the current study, joint interference effects are known to occur even if the participants are not interacting and might in fact be seated in different rooms (see, for example, Experiment 3 in Atmaca, Sebanz, & Knoblich, 2011). In other words, the belief of acting with another person can be sufficient to induce representations of that person's task in the participant. If speakers similarly represent an absent partner's utterances (even if the partner's speech is not perceived but only imagined), and if they do so in a way that is similar to how they represent their own utterances, then they should be affected by these representations while producing their own utterances.

In this study, we tested two speakers simultaneously, but they were seated in separate rooms and could not hear each other. However, they could infer what their partner was about to say because they could see their partner's instructions as well as their own (i.e., they held beliefs about their partner's task). Therefore, we tested whether representations of others' utterances are formed in a non-interactive joint task setting, in which two individuals produce language alongside each other, but are not using language to communicate. A further question we investigated is whether representations of others' utterances are content-specific – whether people represent *what* their partner is saying, or whether they simply represent that their partner is naming or, more generically, that their partner is acting in some way. This question is motivated by a recent debate in the literature on joint interference effects, as we explain in more detail in the following section.

Representing Another's Task

A clear demonstration that one's partner's task is automatically represented comes from experiments that used the Simon effect (Hommel, Colzato, & van den Wildenberg, 2009; Sebanz, Knoblich, & Prinz, 2003, 2005; Sebanz, Knoblich, Prinz, & Wascher, 2006; Tsai, Kuo, Jing, Hung, & Tzeng, 2006; Tsai, Kup, Hung, & Tzeng, 2008; Vlainic, Liepelt, Colzato, Prinz, & Hommel, 2010). In the classic (*solo*) Simon task, participants respond to non-spatial properties of visual stimuli (e.g., color) with their right or left hand. The spatial location or orientation of the stimulus, though irrelevant to the task, affects performance. In particular, a SR spatial compatibility effect is observed: Responses to left stimuli with the right hand and to right stimuli with the left hand (incongruent SR mapping) are slower than responses to left stimuli with the left hand and responses to right stimuli with the right hand (congruent SR mapping).

Importantly, this effect is absent (e.g., Sebanz, et al., 2003), or greatly reduced (e.g., Sebanz, et al., 2005), in a go/no-go version of the same task. However, when the task is distributed across two people, so that each participant performs only half of the task (as in the go/no-go version) but alongside a partner who takes turns with them and performs the other half of the task, a compatibility effect is often observed (this is termed the *joint* Simon effect). Studies have also demonstrated joint effects in other tasks, such as the SNARC task (Atmaca, Sebanz, Prinz, & Knoblich, 2008) and the Flanker task (Atmaca, et al., 2011).

What is the source of joint interference effects? According to the *task co-representation* account, the partner's task (set of SR mappings) is automatically represented in a joint-task setting. The irrelevant features of the stimulus activate the partner's response, which then interferes with the participant's own response (Sebanz, et al., 2005). Specifically, this account predicts that interference occurs when the partner's response and the participant's own response are incongruent or incompatible.

So, for example, a participant sitting on the right of the screen responded to the color of a stimulus (e.g., a photograph of a human hand) more slowly if the stimulus evoked the response of her partner sitting on the left of the screen (i.e., the hand was pointing to the left; see Sebanz et al., 2003) than if the stimulus evoked her own response (i.e., the hand was pointing to the right). According to the task co-representation account, this occurred because she coded her partner's response as a "left" response, and she coded her own response as a "right" response. The orientation of the stimulus evoked these spatial codes, as did preparing one's own response. If the codes mismatched, interference occurred: Execution of the response was slowed down.

Note, however, that the task co-representation account has been challenged. According to the *actor co-representation* account (Wenke et al., 2011), people represent whether another actor might (potentially) be responding on the current trial, but they do not represent another actor's task, and therefore they do not represent another actor's response. According to the actor co-representation account, joint interference effects are thus due to an actor identification conflict rather than to response conflict. In other words, it is the fact that one's partner might (potentially) be responding on the current trial that matters, not the nature of their response. The account predicts that double-response trials (i.e., trials that evoke a response from both co-actors) should show longer latencies than single-response trials (i.e., trials that evoke a response from only one co-actor). This pattern should occur irrespective of whether a congruent response or an incongruent response is evoked.

Philipp and Prinz's (2010) results speak in favor of the actor co-representation account. They reported a joint actor-face compatibility effect in an experiment in which speakers had to respond to colored shapes by uttering their own or their partner's name. Together with the target shape, they were shown a picture of their own or their partner's face (as task irrelevant distractors). Responses were faster when participants saw their own face than when they saw their partner's face, regardless of which name they used. Participants in this study appear to have interpreted the pictured face as a cue to whose turn it was to respond (see Wenke et al., 2011, p. 165).

Interestingly, various studies found joint interference effects when partners were seated in the same room but could not perceive each other (Sebanz, et al., 2005; Vlainic, et al., 2010). Such effects occurred even when participants sat in a room on their own but were led to believe another person was performing the task with them. The finding occurred both when the participant obtained feedback while carrying out the task (i.e., a stimulus on the participant's screen signalled when the believed partner produced a response; see Ruys & Aarts, 2010; Tsai, et al., 2008) and when no feedback was available (Atmaca, et al., 2011).

The literature on joint tasks has focused on tasks involving manual responses. However, a few studies have shown joint interference effects with verbal responses. In addition to Philipp and Prinz's (2010) study reported above, Pickering and McLean (2013) had participants perform a joint version of the Stroop task, in which each participant responded to only one color. Stroop interference was larger in this joint version than in the go/no-go version, where only one participant responded to only one color, and the other did not respond (at least when the participants in the joint version of the task provided feedback to each other's responses). This showed that participants represented their partner's potential response and that this caused additional interference with their own response on incongruent trials.

The present study investigated whether joint task effects can occur when participants are asked to produce the names of pictured entities. Unlike Philipp and Prinz (2010) and Pickering and McLean (2013), the response set was large and stimulus-response mappings were not arbitrary. Participants saw pairs of pictures, and were cued to name one picture or both pictures in a particular order. They were also cued about whether their partner simultaneously named the picture(s) either in the same ("congruent") way or in a different ("incongruent") way, or did not name any pictures. We investigated whether belief about what one's partner was doing affected naming latencies.

Note that our participants could not interact: They named pictures at the same time as one another, but could not hear each other. Therefore, participants might not represent their partner's utterances at all. The studies that showed that non-interacting participants display joint interference effects (e.g., Atmaca, et al., 2011) all used manual responses. We do not know whether the same would hold for verbal responses, particularly because language is more tightly associated with communicative situations than manual actions. If other-representations are not formed, we would expect naming latencies to be unaffected by beliefs about one's partner's task. We term this the *no co-representation* account.

Our study attempted to test among the task co-representation account, actor co-representation, and no co-representation account. The task co-representation account assumes that participants represent the content of their partner's response; therefore, in the case of verbal responses, this account predicts that participants should activate lexical representations corresponding to their partner's responses. The actor co-representation account predicts that participants represent that their partner is responding but not the content of her response. The no co-representation account predicts that participants do not represent their partner's behavior.

In four joint picture-naming experiments, pairs of participants sat in separate soundproof rooms, received instructions that informed them of their task and their partner's task, and were presented with pairs of pictures. In Experiment 1, participants named both pictures and we manipulated whether they believed their partner was naming the pictures in the same order, the reverse order, or did not name the pictures. Experiment 2 replaced the final condition with a condition in which participants believed their partner was deciding whether the pictures were from the same category or not. In Experiment 3, participants named one picture and believed their partner named the same picture, the other picture, or no picture. Experiment 4 was the same as Experiment 3 except that the pictures were degraded.

Both the task and the actor co-representation accounts predict that participants should take longer to name pictures when they believe their partner is also naming pictures than otherwise. In addition, the task co-representation account predicts that participants should take longer to name pictures when they believe that their partner is naming pictures in an incongruent way (either a different picture or the same two pictures in the opposite order) than a congruent way. The actor co-representation account predicts that participants will be unaffected by congruency (i.e., they would represent *that* their partner is responding, but not *what* she is saying). Finally, the no co-representation account predicts that naming will be unaffected by participants' beliefs about what their partner is doing.

In all experiments, we also varied whether the two pictures were semantically related or not. Semantic relatedness leads to interference in picture-word tasks (Schriefers, Meyer, & Levelt, 1990). But in picture-picture tasks when participants name one picture, studies have found no semantic effects (Damian & Bowers, 2003; Navarrete & Costa, 2005) or facilitation (La Heij, Heikoop, Akerboom, & Bloem, 2003). In contrast, people are slower to initiate naming pairs of semantically related than unrelated pictures (Aristei, Zwitserlood, & Abdel Rahman, 2012; see also Smith & Wheeldon, 2004). Therefore we predicted semantic inference when participants named both pictures (Experiment 1 and 2) but not when they named one picture

(Experiment 3 and 4), regardless of their beliefs about their partner's behavior. It is possible that semantic interference would be affected by beliefs about one's partner's task, if the task co-representation account is correct, and we address this in relation to individual experiments below.

Finally, note that we manipulated participants' beliefs about their partner's task while holding partner's presence constant, as the mere presence of another person affects performance in verbal tasks (see e.g., Klauer, Herfordt, & Voss, 2008). To this aim, the partner's instructions were displayed on the screen together with the participants' instructions, and participants attended to the same stimuli as their partner. In this way, it was also possible to manipulate participants' beliefs on a trial-by-trial basis, and thus investigate whether representations of others' utterances are continuously updated.

Experiment 1

In Experiment 1, two differently colored pictures were presented simultaneously to both participants. On each trial, before the pictures appeared, an instruction screen showed the names of the two participants accompanied by the words *red*, *blue,* or *no*. *Red* and *blue* corresponded to "go" trials: The participant was instructed to name the picture presented in the given color first, and then name the other picture. *No* corresponded to "no-go" trials: The participant was instructed to give no response.

We manipulated the order in which the other participant (the partner) concurrently named the pictures (Partner's task), as follows. On trials on which the two participants were assigned the same color (*blue-blue* or *red-red*), they named the pictures in the same order, therefore producing the same verbal response (SAME condition). On trials on which the two participants were assigned different colors (*blue-red* or *red-blue*), they named the pictures in reverse order, therefore producing different verbal responses (DIFFERENT condition). Finally, when either of the participants was assigned a “no-go” trial (*red-no*, *blue-no*, *no-red*, *no-blue*), one participant named the pictures while their partner produced no response (NO condition). See Figure 1 for a summary of the Partner’s task manipulation employed in Experiment 1.

INSERT FIGURE 1 ABOUT HERE

We expected Partner’s task to affect naming latencies. Specifically, according to the task co-representation account, naming latencies should be longer in the DIFFERENT condition than in the SAME condition. This is because it assumes that other-representations are content-specific; that is, they specify the lexical items that the partner is retrieving. Note that, because the speakers always named both pictures, their utterances always contained the same lexical items. However, when the order differed, the picture that the speaker had to name second was the picture that their partner had to name first. Therefore, in the DIFFERENT condition the representation of the partner’s response might enhance activation of the second picture’s name. This would in turn result in the second picture’s name competing more for selection with the first picture’s name, leading to longer naming latencies. Instead, when the representation of one’s partner response specified the same order as the representation of one’s own response (SAME condition), the first picture’s name was also the word that one’s partner had to retrieve first. Therefore, the activation level of the first picture’s name might be raised and competition with the second picture’s name might be reduced (compared to the DIFFERENT condition). The predictions of the task co-representation account are summarized in Figure 2 (panel A).

Alternatively, according to the actor co-representation account, speakers do not represent the content of their partner's response, but they represent whether their partner is responding on the current trial or not. If this is the case, the relationship between self- and other-representations would not affect processing, and hence naming latencies would be equivalent in the SAME and DIFFERENT conditions. However, naming latencies should be longer in the SAME and DIFFERENT conditions than in the NO condition, because in the latter participants believe that their partner is not responding. The predictions of the actor co-representation account are summarized in Figure 2 (panel B).

Finally, according to the no co-representation account, another person's utterances should not be represented at all under the conditions tested in our experiment (i.e., in the absence of interaction). The account therefore predicts that the Partner's task manipulation will have no effect (i.e., there will be no difference between the SAME, DIFFERENT, and NO conditions). This scenario is presented in Figure 2 (panel C).

As an additional manipulation, participants saw either two semantically related (e.g., *apple – banana*) or two semantically unrelated pictures (e.g., *apple – blouse*). Smith and Wheeldon (2004) showed that the time to initiate descriptions of moving displays containing two semantically related pictures is longer than the time to initiate comparable descriptions of displays containing unrelated pictures. Furthermore, Aristei, Zwitserlood, and Abdel Rahman (2012) found that German speakers took longer to initiate uttering the names of two pictures (forming a novel noun-noun compound) when they were semantically related than unrelated. Therefore, we expected participants to take longer to name pairs of related pictures than pairs of unrelated pictures (a main effect of semantic relatedness). The manipulation of semantic relatedness therefore served as a manipulation check. In addition, it is possible that Relatedness might interact with Partner's task. If the task co-representation account is correct and other-representations are content-specific, the semantic interference effect could be enhanced in the DIFFERENT condition (compared to the SAME condition) because in the DIFFERENT condition the second picture's name receives additional activation from a representation of one's partner response (see above). Note that an interaction of Partner's Task and Relatedness would be not be predicted by the actor co-representation or the no co-representation account.

INSERT FIGURE 2 ABOUT HERE

Method

Participants

Twelve pairs of previously unacquainted participants from the University of Edinburgh student community were paid £6 to participate. All participants reported being native speakers of English with no speaking or reading difficulties.

Materials

Fifty pictures were paired in two different ways to yield 50 picture-picture pairs (25 semantically related, 25 semantically unrelated; see Appendix A). For example, *apple* was paired once with *banana* (related) and once with *blouse* (unrelated). In turn, *banana* was paired once with *apple* (related) and once with *frog* (unrelated). Since one picture was embedded inside the other, half of the pictures were relatively small (about 250x200 pixels), whereas the others were relatively large (about 600x500 pixels). Of these 50 pictures, 11 pairs were taken from Damian and Bowers (2003), 4 pairs from Navarrete and Costa (2005), and 10 pairs were modeled after materials used in the same studies, but modified to avoid phonologically related names (as the two studies were not conducted in English). Damian and Bowers (2003) and Navarrete and Costa (2005) controlled for visual similarity (i.e., related pairs were not visually more similar than the unrelated pairs). When adapting their materials, care was taken to pair pictures that were visually dissimilar to each other. Eight additional pictures were selected from Snodgrass and VanderWart (1980) to be used on practice trials.

Design

We manipulated three factors, all within participants: Partner's task (henceforth, Partner; SAME vs. DIFFERENT vs. NO), Relatedness (unrelated vs. related), and Size (i.e., size of the first named picture: big vs. small). Partner and Relatedness were also manipulated within items, whereas Size was manipulated between-items. An item was defined in terms of the first named picture (so *apple-blouse* and *blouse-apple* counted as different items). Partner refers to the task assigned to the participant's partner: he or she named the pictures in the same order (SAME; e.g., participant: *apple-blouse*, partner: *apple-blouse*), in reverse order (DIFFERENT; e.g., participant: *apple-blouse*, partner: *blouse-apple*), or did not name any pictures (NO; e.g., participant: *apple-blouse*, partner: *""*). Partner varied on a trial-by-trial basis.

Each picture was repeated 16 times across the experiment. In the SAME and DIFFERENT conditions, each picture was presented four times (twice in a related pair, twice in the unrelated pair). In the NO condition, each picture was presented 8 times (four times in a related and four times in an unrelated pair), in order to get the same number of data as in the other two conditions. Each participant named each picture in first position 6 times, once per cell of the design.

Because exactly the same instructions and stimuli were presented to both participants in the two rooms simultaneously (see below), participants were always correctly informed about the task their partner was concurrently performing. This means that when one participant completed a trial in the NO condition, the other participant did remain silent, and both participants in a pair were assigned to no-go trials equally often.

There were 400 trials in total. These were presented in 4 blocks of 100 trials. Each block comprised an equal number of trials in each condition for both participants. Because the number of pictures could not be divided by four, and because of the requirement that participants named big and small pictures equally often in first position, it was not possible to ensure that each picture was named an equal number of times in each block. However, the order of presentation was pseudo-randomized, separately for each pair and for each block, with the constraint that the same picture never appeared on two consecutive trials. The order of blocks was also counterbalanced across pairs.

On every trial, one picture was red and the other was blue. Participants were cued to start either with the red or the blue picture. To prevent response strategies, we counterbalanced the following factors within each block: color-participant pairing (whether a given participant named the red or the blue picture in first position), color-size (whether the red picture was small and the blue picture was large or vice versa), and order of instructions (whether the color cue for a given participant was displayed in the top half or in the bottom half of the screen).

Procedure

Participants were tested in adjacent soundproof rooms. Each was seated at a distance of about 90 cm in front of a 48-cm 60 Hz LCD monitor; both monitors were connected to the same machine in the control room (so stimulus presentation was simultaneous). Stimulus presentation was controlled using E-Prime (Version 2.0). There was a window between the two rooms, but participants could only perceive each other peripherally when facing the monitors. The two rooms were linked via a communication system that allowed the experimenter to control whether the two participants could hear each other or not. Participants wore headphones through which they could hear their own voice and spoke into high-quality directional microphones (AKG Acoustics, Vienna, Austria, www.akg.com).

Upon entering the lab, the participants were introduced to one another and were each taken to a different, randomly assigned room. The experimenter showed all the (practice and experimental) pictures on the computer screen once, one at a time, with the corresponding names, and asked each participant to repeat them to aid memorization. Immediately afterwards, the pictures were shown again (without names), and each participant was asked to produce the correct name. Participants completed this phase in parallel, but without hearing one another. The experimenter listened to both participants and provided correct feedback in case either of them made a mistake or could not retrieve the name. Participants were then informed that they would “work together” and were called out of the booths. Instructions were delivered to both participants at the same time in the control room. The instructions stressed that they should try to name both pictures as quickly as possible, while still preserving accuracy and clear pronunciation.

The participants then went back to their respective booths and performed 20 practice trials. These were similar to experimental trials but used only practice pictures, which were matched to form semantically unrelated pairs. They were presented in a different random order for each pair. For each participant, on four of the practice trials the partner named the pictures in the same order; the partner named the pictures in reversed order on four other trials, and in six trials the partner named no picture. In the remaining six trials, the participant remained silent while the partner named the pictures. Finally, participants began the experimental phase.

On each trial, first a fixation cross was displayed for 1000 ms, then a display (2000 ms) that showed the participants names, each followed by an instruction word. After a 500-ms blank, two pictures (one red, one blue) were displayed simultaneously (for 400 ms). Each trial was concluded by a 1500 ms inter-stimulus interval. We reasoned that the speed of picture presentation would make it difficult for participants to use color information strategically (i.e., delaying processing the second picture until they have begun uttering the first picture’s name), and thus would encourage them to carry out a substantial amount of processing of both pictures in parallel before speech onset (Wagner, Jescheniak, & Schriefers, 2010).

The 4 blocks were separated by breaks of variable length. The participants were allowed to rest for as long as they required. The experimenter checked with both participants that they were happy to continue before resuming the experiment. An experimental session lasted about 1 hour.

Recording and Data Analysis

A 75-ms beep (inaudible to participants) was used to mark stimulus presentation and was recorded together with the participants’ responses (on three separate channels, sampling rate: 48000 Hz) via an M-Audio FireWire 1814 device (inMusic, Cumberland, RI, www.m-audio.com) in Adobe Audition (Version 4.0). Beep onsets were automatically tagged using Audacity (Version 1.2.5). Recordings were then pre-processed to reduce background noise. Speech onsets were first automatically tagged using the Silence finder algorithm in Audacity and later checked manually for lip smacks and other non-speech noises. Naming latencies were defined as the time from beep onset to response onset.

We analyzed the data with Generalized Linear mixed-effects models (Bayeen, Davidson, & Bates, 2008), as implemented in the lme4 package (Bates, Maechler, & Dai, 2008; Version 0.999375-28) in R (Version 2.13.1). We chose this method because it allows both naming latency and accuracy data to be analyzed within the same statistical framework. The LME analyses reported below used a normal link function and modeled untransformed naming latencies. However, because the distribution of naming latencies was noticeably right-skewed, we also conducted LME analyses on inverse-transformed naming latencies (see Appendix B). For the accuracy data we used a logistic link function (Jaeger, 2008); a correct response (see Results below) was coded as 0, whereas an incorrect response was coded as 1.

In the LME analyses, we first focused on selection of the random effects structure (keeping the fixed effects structure maximal). We started with the full random effect structure, including random slopes (for all factors and their interaction) and random intercepts for both subjects and items. Since random slopes are only appropriate for within factors, we included by-subjects random slopes for Partner, Relatedness, and Size, and by-items random slopes for Partner and Relatedness. If the model with full random structure did not converge, we simplified it by removing higher order terms; we started with the three-way interaction by-subjects and proceeded with removing the two-way interactions, one at a time, first by subjects, then by items. Once we found a model that converged, we used backward selection to select the slopes that contributed to model fit (with the alpha-level set to .1 to account for the conservativity of such tests).[1]

Once we identified the random effects structure as described above, we proceeded to select the best-fitting fixed effects structure. We started by fitting the complete model (including the main effects of Partner, Relatedness, and Size, the three two-way interactions, and the three-way interaction); we then removed predictors that were not significant from the model, using a backward stepwise procedure, and stopped whenever removing a predictor caused a significant loss of fit (assessed using a log-likelihood ratio test, with the alpha-level set to .05). Marginal fixed effects ($.05 < p < .1$) are reported below for completeness, but we chose not to include them in the best-fitting models reported in Appendix B, and we do not attempt to interpret them. Note that in case no fixed effects contribute significantly to model fit, the best-fitting model is the intercept-only model; such models are not reported.

All predictors were contrast-coded. For Partner, we defined two planned contrasts: *naming vs. no* compared the DIFFERENT and SAME conditions against the NO condition; *same vs. different* compared the SAME against the DIFFERENT

---

[1] Barr et al. (2013) recently argued in favour of keeping random effect structures maximal in LME analyses. Partly because of the complexity of our experimental design, models with full random effect structure seldom converged, despite fixing some of the parameters (random correlations between intercepts and slopes) to zero. In Appendix B, we report LME analyses that used the maximal random effect structure that converged (for a similar approach to reporting LME analyses, see Jaeger, Furth, & Hilliard, 2012), and we report significant coefficients from the complete model (i.e., the model containing all fixed effects and interactions).

condition. Specifically, the following coding scheme was used for all analyses reported below, unless otherwise stated: Partner1: SAME: 1/3, DIFFERENT: 1/3, NO: -2/3; Partner2: SAME: 1/2, DIFFERENT: -1/2, NO: 0; Relatedness: related: 1/2, unrelated: -1/2; Size: small: 1/2, big: -1/2.

Results

We only report descriptive statistics and the results of likelihood ratio tests in this section. See Appendix B for full reports of the best fitting models corresponding to the analyses reported in this section (and in the Results sections of the other experiments reported in this paper).

Accuracy

Responses were coded as correct and entered into the onset time analysis (see below) only if both pictures were named correctly. Incorrect responses fell into 4 different categories: naming errors (the wrong name was used); disfluency (the correct name was used, but the response contained hesitations or repetitions); order error (the second picture was named before the first picture); no response (the participant remained silent when he or she had to respond). Error counts and percentages are reported in Table 1.

INSERT TABLE 1 ABOUT HERE

Removing by-items random slopes for the factor Relatedness significantly harmed fit (χ^2 (4) =19.70, p<.001). Interestingly, participants produced more incorrect responses when their partner named than when he or she remained silent (Partner 1: log-odds B= .24, SE = .11, z = 2.23), and also fewer incorrect responses in the SAME than in the DIFFERENT condition (Partner 2: log-odds B = -.23, SE =.08, z= -2.75; see Table 1B). Accordingly, Partner contributed significantly to model fit (χ^2 (2) =13.10, p<.01). The main effect of Relatedness was only marginally significant (χ^2 (1) =3.54, p= .06) and the corresponding coefficient is not reported in Table 1B.

Naming Latencies

Naming latencies longer than 3000 ms or shorter than 300 ms were considered outliers and removed from all analyses. However, there were no such cases in Experiment 1. Then by-participant means and standard deviations were computed. Values that were more than 3 standard deviations from the by-participant mean (1.5%) were replaced with the cut-off value.[2] Mean latencies are reported in Table 2; see also Figure 3, Panel A.

INSERT TABLE 2 ABOUT HERE

INSERT FIGURE 3 ABOUT HERE

No slopes contributed to fit, so only random intercepts were retained. The main effects of Partner (χ^2 (2) =7.80, p<.05) and Relatedness (χ^2 (1) =11.32, p<.001) contributed to fit. The three-way interaction between Partner, Relatedness, and Size was only marginally significant (χ^2 (2) =5.15, p=.08) and is not included in Table 2B. All two-way interactions, including the interaction of Partner and Relatedness were not significant.

Participants took significantly longer to start speaking when their partner was also preparing to speak (naming: 876 ms) than when their partner remained silent

[2] Additional analyses performed on the complete data set (with only the outliers above 3000ms or below 300ms removed) yielded a similar pattern of results.

(NO: 864 ms; Partner1: B = 14 ms, SE = 5, t = 2.79). However, average naming latencies were very similar in the DIFFERENT (875 ms) and SAME (877 ms) conditions (Partner 2: B = 1 ms, SE = 4, t = .17; see Table 2B). Therefore, the main prediction of the task co-representation account (i.e., that latencies would be longer in the DIFFERENT than in the SAME condition) was not confirmed. In addition, this finding is inconsistent with the no co-representation account, but fully consistent with the actor co-representation account.

Finally, replicating previous findings (e.g., Aristei et al., 2012), participants took longer when the two pictures were semantically related than when they were unrelated (880 vs. 864 ms). Interestingly, the semantic interference effect was similar across conditions (12 ms in DIFFERENT, 17 ms in SAME, 17 ms in NO), as demonstrated by the lack of an interaction between Partner and Relatedness.

Discussion

Experiment 1 showed that participants took longer to initiate a naming response when they believed their partner was also preparing a naming response. This finding rules out the no co-representation account (Figure 2, Panel C).

The results of Experiment 1 do not fully support the task co-representation account (Figure 2, Panel A). Participants made more errors when their partner was preparing an incongruent (DIFFERENT) than a congruent (SAME) response. This finding might suggest that speakers experienced more interference when they believed their partner was preparing an incongruent response, as predicted by the task co-representation account. However, this pattern in naming accuracy was not confirmed by a similar pattern in naming latencies. Participants were no slower when they believed their partner was preparing an incongruent response than when they believed they were preparing a congruent response. Additionally, the semantic interference effect (longer latencies for related than unrelated responses) was no greater in the DIFFERENT (12 ms) than in the SAME condition (17 ms). On the contrary, the results of Experiment 1 are fully consistent with the actor co-representation account (Figure 2, Panel B), which claims that people represent whether a response occurs or not, but they do not represent the content of the response itself. Consistently, latencies were longer when both participants named pictures (i.e., DIFFERENT and SAME conditions) than when only one did (i.e., NO condition), regardless of the relationship between the participant and their partner's response. However, we must consider alternative explanations of this finding.

The conditions in which participants responded slowly (SAME and DIFFERENT) are the ones in which two "go" instructions are displayed on the screen. It might be that, despite being addressed by their first name (a highly salient stimulus; e.g., Wood & Cowan, 1995), participants were distracted by the presence of their partner's instruction more when it was a "go" instruction than when it was a "no-go" instruction. "Go" instructions are words of the same type (color names), whereas "no-go" instructions used a clearly different word ("no"). Therefore, "go" instructions are more similar to each other than they are to "no-go" instructions. Similarity might cause interference between memory representations for one's own and the partner's instructions.

Note that participants rarely performed their partner's task by mistake, which seems to suggest that they rarely misremembered the instructions. There is however some indication that this occurred more often in the DIFFERENT condition (on 2.3% of trials speakers named the pictures in their partner's order), compared to the NO condition (on 1.2% of trials speakers gave no response).[3] But more importantly, the similarity explanation cannot account for why latencies were no longer in the DIFFERENT than in the SAME condition (where instructions were identical, so interference between memory representations is unlikely to have occurred).[4] We return to this issue in the Discussion to Experiment 2, where we replaced the "no-go" instructions with (a different kind of) "go" instructions. Here we note that interference between memory representations for the instructions cannot account for the finding that naming latencies were slower in both naming conditions (DIFFERENT and SAME) than in the NO condition.

Therefore, we conclude that participants represented whether it was their partner's turn to respond on any given trial and that they experienced interference whenever both they and their partner were preparing a response concurrently. But what sort of mechanism could be responsible for this interference effect? Participants might represent that their partner was "doing something" at the same time that they prepared their response. If this is the case, we expect that a belief that one's partner is performing *any* task would slow down the process of naming to the same extent as a belief that they are naming pictures.

---

[3] Of course, it was not possible to make a mistake of this type in the SAME condition, as both participants were given the same instruction.

[4] In addition, in another study (Van de Cavey, Gambi, MacKenzie, Nelissen, & Pickering, 2012, September) we conducted a related experiment where participants described simple scenes using active or passive sentences. In that study, we replaced the word "no" with "grey" (keeping instructions the same; i.e., to remain silent) and found that latencies were still longer when participants believed their partner was speaking than when they believed their partner was silent.

Alternatively, the interference effect could arise because the same mechanisms (i.e., language production mechanisms) are used to represent one's partner naming response and to prepare one's own naming response. If so, we predict less interference when one's partner is performing a different (non-naming) task than when one's partner is preparing a naming response. Experiment 2 was designed to decide between these alternatives.

Experiment 2

In Experiment 2 we replaced the NO condition with a semantic categorization (hence, CAT) condition. The SAME and DIFFERENT conditions were the same as in Experiment 1. In the CAT condition, participants named the pictures while their partner judged whether the two pictures were from the same semantic category (as discussed below). For example, *goat* and *pig* were considered as belonging to the same semantic category (animal), while *goat* and *cup* were not. Responses to the categorization task were given by speaking "yes" or "no" into the microphone.

Therefore, all trials required an overt verbal response from both participants. Consequently, if imagining one's partner performing *any* task was driving the effect we observed in Experiment 1, we should find no difference between the SAME, DIFFERENT, and CAT conditions. We chose the CAT task because it would be particularly similar to the naming task and therefore likely to share the same pool of attentional resources. First, both tasks involve visual processing of the pictures and retrieving the concepts associated with the depicted entities. In addition, both tasks involve the articulation of an overt verbal response. Crucially, however, the tasks differ in the extent to which they engage language production mechanisms (i.e., lexical retrieval).

In the naming task, speakers clearly need to access the lexical items corresponding to the concepts depicted in the pictures. In the categorization task, however, speakers need not access the lexical items, as the task can be performed on the basis of conceptual information alone. That indeed speakers do not access the names of pictures when they do not perform a naming task is confirmed by Jescheniak and Levelt (1994, Experiment 2), in which participants decided whether a picture matched a previously presented word or not. Despite the fact that participants were exposed to the picture names on matching trials, there was no evidence that they (re-)accessed the picture names at the moment of performing the task, as shown by the lack of a frequency effect.

More generally, it is common practice in psycholinguistics to contrast picture categorization with picture naming in order to disentangle the contribution of perceptual and conceptual processing from lexical retrieval processes proper (for a similar logic, see for example Almeida, Knobel, Finkbeiner, & Caramazza, 2007 and references therein). Therefore, we assume that the naming task engages language production mechanisms (i.e., lexical retrieval) to a greater extent than the categorization task. If the interference effect in Experiment 1 was specific to representing that one's partner is preparing to name (i.e., is engaging in lexical retrieval) it should be replicated in Experiment 2.

Finally, we retained both the SAME and the DIFFERENT conditions from Experiment 1 to provide another test of their comparative effects. The accuracy data in Experiment 1 seemed to suggest that the DIFFERENT condition might induce more interference than the SAME condition. Therefore, we wanted to check whether this effect would be replicated in Experiment 2. The semantic relatedness manipulation was also retained in this experiment (as it ensured that positive and negative responses were balanced in the semantic categorization task).

Method

Participants

Sixteen further pairs of participants from the same community as the pairs in Experiment 1 were paid to participate.

Materials and Design

These were the same as in Experiment 1, except that the NO condition was replaced with the CAT condition. In order for participants to practice the semantic categorization task, we replaced two of the original practice pictures with two new pictures from Snodgrass and VanderWart (1980), so that it would be possible to form semantically related pairs (this also involved re-pairing the original pictures). Various semantic categories were represented in the materials from Experiment 1 (see Appendix A).

Procedure

For the semantic categorization task, participants were instructed to respond to the word *question* (which replaced the word *no*) by answering the following question: “Are the two pictures from the same category?” They responded by saying “yes” into the microphone if they thought the answer was positive, or “no” if they thought the answer was negative. The experimenter provided two examples to clarify what it meant for two pictures to be “from the same category” (one example mentioned *dog* and *snake* as requiring a positive answer, *dog* and *lemon* as requiring a negative answer; the second example mentioned *pen* and *ruler*, as requiring a positive answer, *pen* and *door* as requiring a negative answer). The experimenter also mentioned the relevant superordinate category (i.e., animal; stationery) while illustrating the examples. Otherwise, the instructions and procedure were the same as in Experiment 1.

Recording and Data Analyses

These were exactly the same as in Experiment 1 with regard to naming responses. Responses to the semantic categorization task were also analyzed; latencies for both positive and negative answers were extracted automatically using the Silence Finder feature in Audacity (without manual correction).

Results

Semantic Categorization Task

Four types of responses were coded as incorrect: disfluencies (hesitations, repetitions), wrong responses, missing responses, and task errors (i.e., when participants performed the naming task instead of the categorization task). Overall, participants were highly accurate: they responded correctly on 94.7% of the unrelated trials and on 93.6% of the related trials. There was no significant difference in the number of errors between related and unrelated trials. Task errors amounted to 2.3% of the trials (N=77). We also looked at the latency to respond on correct trials. Based on the overall distribution of responses, we removed the 4 responses that were shorter than 250 ms or longer than 2500 ms. Again, there was no difference between related ("yes") trials (M= 936 ms, SD = 289 ms) and unrelated ("no") trials (M = 944 ms, SD = 291 ms).

Naming Task

Accuracy. Incorrect responses were coded as in Experiment 1, except that another type of error was possible; occasionally (on 2 trials in DIFFERENT, 4 in CAT and 5 in SAME) participants performed the categorization task instead of the naming task. Counts and percentages are reported in Table 3. Removing by-items random slopes for the factor Relatedness significantly harmed fit (χ^2 (4) =18.38, p<.005). The main effect of Relatedness was only marginally significant (χ^2 (1) =2.98, p=.08). No other factor gave a significant contribution to model fit; therefore, the best-fitting model was the intercept-only model (hence this model is not reported in Appendix B).

INSERT TABLE 3 ABOUT HERE

Naming Latencies. Naming latencies longer than 3000 or shorter than 300 ms were considered outliers and removed from all analyses. There were only 2 such cases in Experiment 2. Then by-participant means and standard deviations were computed. Values than were more than 3 standard deviations from the by-participant mean (1.7%) were replaced with the cut-off value.[5]

INSERT TABLE 4 ABOUT HERE

By-participants random slopes for Size (χ^2 (5) =46.16, p<.001) and by-items random slopes for Relatedness (χ^2 (4) =8.21, p<.10) were included. Crucially, Partner contributed significantly to model fit (χ^2 (2) =6.54, p<.05; see Table 3B). Relatedness also contributed significantly to model fit (χ^2 (1) =11.04, p<.001). All two-way interactions and the three-way interaction were not significant.

---

[5] Additional analyses performed on the complete data set (with only the outliers above 3000ms or below 300ms removed) yielded a similar pattern of results.

Participants took longer to start naming when their partner was also preparing to name (891 ms) than when their partner was preparing to categorize the pictures (880 ms; Partner 1: $B = 12$ ms, $SE = 5$, $t = 2.47$). However, average naming latencies were no longer in the DIFFERENT (889 ms) than in the SAME (893 ms) condition (Partner 2: $B = 3$ ms, $SE = 4$, $t = .70$; see Table 3B).

Finally, as in Experiment 1, participants took longer when the two pictures were semantically related than when they were unrelated (897 vs. 878 ms) and the semantic interference effect was similar across conditions (17 ms in DIFFERENT, 28 ms in SAME, 11 ms in CAT), as demonstrated by the lack of an interaction between Partner and Relatedness (see Table 4 and Figure 3, Panel B).

Discussion

The results of Experiment 2 are again not consistent with the no co-representation account, as Partner's task affected naming latencies. In addition, they are not consistent with the task co-representation account. As in Experiment 1, naming latencies did not differ in the DIFFERENT and SAME condition. In addition, and unlike in Experiment 1, the likelihood of producing an incorrect response also did not differ in the two conditions (and did not differ significantly from the CAT condition, either). Therefore, the task co-representation account's prediction that participants experience more interference when they believe their partner is naming the pictures in reverse order than when they believe they are naming in the same order was not supported by the results of Experiment 2. In addition, as in Experiment 1, we found that speakers were slower at naming pairs of semantically related than unrelated pictures, but the semantic interference effect was not larger in the DIFFERENT than in the SAME condition.

Most importantly, naming latencies were longer when speakers believed that their partner was also naming a picture than when they believed that their partner was performing a semantic categorization task. Given that the two tasks share all processing stages except lexical retrieval, we conclude that the process of naming pictures is slowed down by the belief that another speaker is concurrently retrieving the pictures' names. This is not consistent with a version of the actor co-representation account in which speakers only represent whether it their partner's turn to respond on the current trial. Rather, the results of Experiment 2 suggest that speakers specifically represent whether another speaker is planning to name, and not just whether their partner is about to respond (in any way). We return to this issue in the General Discussion.

In addition, we note that naming latencies were longer in the Partner-naming than in the CAT conditions, even though the CAT condition was also associated with "go" instructions for the participants' partner. This weakens the concern raised in the Discussion to Experiment 1 that interference on Partner-naming trials in that experiment was due to the presence of similar instructions (both "go" instructions) and not to a representation of the partner's response.

However, it is possible that the finding that participants represent whether their partner is about to name a picture or not is specific to the situation in which participants have to encode both pictures' names, while simultaneously formulating an utterance plan that specifies order of mention. These requirements might make the task rather demanding and perhaps more sensitive to interference from a representation of the other person's response. To investigate whether similar effects would occur when speakers were naming single words, we conducted Experiment 3.

Another aim of Experiment 3 was to test whether representations of others' responses are formed when those responses bear no relationship to one's own responses. In Experiment 1 and 2 participants might have formed representations of their partner's responses because those responses always overlapped in content with their own responses, and were therefore perceived as somehow associated with their own responses. In Experiment 3, we tested a condition in which the partner produced a response that was completely different from the participant's concurrent response (except for the fact that the corresponding visual stimulus was co-located with the stimulus the participants responded to).

Finally, in Experiment 3 we provided yet another test of the task co-representation account. Proponents of this account might note that in Experiments 1 and 2 the content of the partner's response (i.e., the identity of the retrieved lexical items) was in fact identical in the SAME and DIFFERENT conditions, as the only difference was in the order of mention. It is conceivable that people do indeed represent the content of their partners' responses, but not order. Therefore in Experiment 3 we changed the instructions so that the partner named either the same picture or a different picture than the participant.

Experiment 3

Experiment 3 was similar to Experiment 1, except that participants named one picture rather than both pictures. Participants named the picture in the assigned color, and ignored the other picture. Therefore, we included a condition in which participants named a picture while their partner remained silent (NO), a condition in which participants named the same picture (SAME), and a condition in which participants named different pictures (DIFFERENT). Of course, participants therefore believed their partner was naming one picture or no picture.

We reasoned that the task would be much less demanding than the task in Experiment 1. Speakers had to retrieve and produce only one word. Therefore, we expected them to respond at shorter latencies. Of course, this task also requires speakers to successfully ignore the non-target picture. There is evidence that distractor picture names are accessed during target picture naming (e.g., Meyer & Damian, 2007; Morsella & Miozzo, 2002; Navarrete & Costa, 2005). However, distractor picture names should be less activated than in Experiments 1 and 2. This is because in those experiments participants retrieved and articulated both pictures' names, while in the current experiment they were explicitly instructed to ignore the distractor picture. If the reduced demands of this task make it somewhat more impenetrable to interference, then the effect of Partner's task might be reduced in Experiment 3 compared to Experiment 1.

In addition, participants in the DIFFERENT condition of Experiment 3 produced a response that did not overlap in lexical content with their partner's response. Therefore, this made it easier for participants to ignore their partner's response in this condition. It is possible that this will further reduce the size of the interference effect (in the DIFFERENT condition only). Nevertheless, the literature on shared representations of manual responses shows that one's partner's responses are represented even when they are irrelevant to one's own task (see Knoblich, et al., 2011). Therefore, it is also possible that participants will form representations of their partner's response even when it is irrelevant to their own task, as in the DIFFERENT condition of Experiment 3.

Finally, Experiment 3 provides an additional test of the task co-representation account. We have assumed that participants naming the same two pictures in opposite orders are performing different tasks, so that their beliefs about what their partner is doing might interfere with their own performance. However, it is conceivable that participants do not represent the order in which their partner names the pictures (and therefore merely represent which pictures their partner names). If this were the case, then the SAME and DIFFERENT conditions in Experiments 1 and 2 would in fact be equivalent in terms of what other-representations are formed. But this is not the case in Experiment 3, as the SAME and DIFFERENT conditions involve distinct target pictures. Therefore the task co-representation account predicts that naming latencies will be longer in the DIFFERENT than the SAME condition. In contrast, the actor co-representation account predicts that naming latencies will be similar in the SAME and the DIFFERENT condition.

Finally, note that we did not expect a main effect of semantic relatedness in this experiment, as most previous studies in which participants named one target picture found no effect of the semantic relatedness of a distractor picture (Damian & Bowers, 2003; Navarrete & Costa, 2005; but see La Heij et al., 2003 for a study reporting semantic facilitation). We retained the manipulation of semantic relatedness in the interest of comparability across experiments.

Method

Participants

Thirteen further pairs of participants from the same community as participants in Experiment 1 and 2 were paid to participate. One pair was removed from the analyses because one of the participants produced exceptionally long naming latencies.

Materials, Design, Procedure, Recording, and Data Analysis

These were the same as in Experiment 1, except that participants were instructed to interpret the color cue as indicating that they had to name the picture presented in that color (and ignore the other picture).

Results

Accuracy

Recordings for 21 (0.2%) trials could not be analyzed due to experimental error or technical problems. For the remaining trials, we coded whether the response was correct or incorrect. Incorrect trials were trials on which the participant used the wrong name for the picture, used the correct name but produced it disfluently, or performed the wrong task (either did not name a picture when they had to, or named the wrong picture; there were 10 such cases in DIFFERENT, 11 in SAME, and 10 in NO).

INSERT TABLE 5 ABOUT HERE

No random slopes contributed to model fit (all *p*'s > .1). As can be seen from Table 5, the likelihood of producing an incorrect response was not affected much either by Partner or by Relatedness. The interaction of Relatedness and Partner contributed to fit only marginally ($\chi2$ (2) =4.73, p=.09); no other factor or interaction approached significance. Therefore, the best-fitting model was the intercept-only model (hence this model is not reported in Appendix B).

Naming Latencies

Naming latencies longer than 3000 or shorter than 300 ms were considered outliers and removed from all analyses. There were 6 such cases in Experiment 3. Values that were more than 3 standard deviations from the by-participant mean (1.6%) were replaced with the cut-off value.[6] Mean latencies are reported in Table 6 (see also Figure 3, Panel C). By-participant random slopes for Size (χ^2 (5) =10.17, p=.07) and by-item slopes for Relatedness (χ^2 (4) =22.10, p<.001) were included. Among fixed factors, only the interaction of Relatedness and Size significantly contributed to fit (χ^2 (1) =4.74, p<.05). The factor Partner was not significant (χ^2 (2) =4.33, p=.11). See Table 5B.

INSERT TABLE 6 ABOUT HERE

In order to resolve the interaction, we used treatment coding to fix one of the two levels of the factor Size to zero and refitted the model shown in Table 5B twice to derive estimates for Relatedness separately for small (coding for Size: big = 1, small = 0; i.e., “small” as baseline level), and big (coding for Size: big = 0, small = 1; i.e., “big” as baseline level) pictures. While there was a non-significant tendency towards semantic facilitation for big pictures (B= -8, SE = 7, t=-1.14), small pictures showed a tendency towards semantic interference, which approached significance (B= 14, SE = 7, t=1.96).

Discussion

[6] Additional analyses performed on the complete data set (with only the outliers above 3000ms or below 300ms removed) yielded a similar pattern of results.

Unlike in Experiments 1 and 2, the effect of Partner was not significant in Experiment 3, although we found a numerical tendency in the same direction as in Experiments 1 and 2. Given that Experiment 3 had the same number of participants and items as Experiment 1, it is possible that the effect of Partner is less apparent under simple naming conditions. This could be because the simple-naming task is particularly undemanding (and thus less susceptible to interference). We explore this possibility further in Experiment 4. Overall, we found little evidence that participants represented their partner's response.

Finally, we found an interaction between Relatedness and Size in the latency analyses. This interaction was driven by a tendency towards greater semantic interference when speakers where naming small pictures versus big pictures. Interestingly, we note that Damian and Bowers (2003), who reported no effect of semantic relatedness in a similar task, asked their participants to name big pictures but not small pictures. Unfortunately, it is not possible to determine the source of this interference effect. Given that our materials were not fully controlled for visual similarity, we cannot rule out that the relatedness effect was in fact caused by visual similarity (that could be greater between related than unrelated picture pairs).

Experiment 4

Experiment 4 was identical to Experiment 3, except that all stimuli were visually degraded. Degradation is known to cause quite large increases in picture naming latencies (e.g., Mädebach, Jescheniak, Oppermann, & Schriefers, 2011), most likely because participants take longer to retrieve the concept associated with the pictures. If interference from a representation of one's partner's task occurs only when the task is particularly demanding, then we should observe an effect of Partner's Task in this Experiment, unlike in Experiment 3.

According to the task co-representation account, participants should take longer to initiate naming in the DIFFERENT than in the SAME condition.[7] The actor co-representation account, instead, would predict slower latencies in the DIFFERENT and SAME conditions than in the NO condition, but no difference between the SAME and the DIFFERENT conditions.

Method

Participants

Twelve further pairs of participants from the same community as participants in Experiment 1, 2, and 3 were paid to participate.

---

[7] Note that Mädebach et al. (2011) demonstrated that distractor pictures are processed more shallowly (as shown by a lack of facilitation from phonologically-related distractor pictures) when the target picture, the distractor picture, or both (like in the present experiment) are visually degraded. They argued that processing constraints limit the amount of resources that can be devoted to encoding the name of distractor pictures in such cases. Because degradation makes it harder to retrieve the concepts associated with distractor pictures, it is therefore possible that participants would be unlikely to represent their partner's utterances in this Experiment. However, this could only be the case for the DIFFERENT condition; if the task co-representation account is correct, participants should still be facilitated in the SAME condition. Whatever the effects of degradation, the task co-representation account predicts longer latencies in the DIFFERENT than in the SAME condition in Experiment 4.

Materials

These were the same as in Experiments 1 to 3, except that a mask of parallel white lines (see Mädebach, et al., 2011) was superimposed on the pictures in order to conceal part of the lines. The proportion of masked lines varied from picture to picture; we tried to keep the proportion of masked lines constant for each picture across the related and unrelated condition, but this was not always possible because the mask superimposed on small pictures partly overlapped with the contours of the big picture. However, as explained above, we were not expecting participants to retrieve the distractor's name in this experiment and, therefore our interest was not focused on the semantic relatedness manipulation.

Design, Procedure, Recording and Data Analysis

These were the same as in Experiment 3.

Results

Accuracy

Recordings for 2 trials could not be analyzed because of experimental error. The remaining trials were coded as correct or incorrect as in Experiment 3. The participants performed the wrong task (either did not name a picture when they had to, or named the wrong picture) on 31 trials in the DIFFERENT condition, 24 in the SAME condition, and 33 in the NO condition. Counts (and percentages) of incorrect trials are given in Table 7, broken down by Partner, Relatedness, and Size of the named picture. By-participant random slopes for Size contributed to model fit ($\chi2$ (5) =35.97, $p<.001$). The interaction of Partner and Size contributed to model fit significantly ($\chi2$ (2) =10.57, $p<.01$). When naming big pictures, speakers made 58 errors in DIFFERENT, 39 in SAME, and 34 in NO. When naming small pictures, they made 48 errors in DIFFERENT, 58 in SMALL and 68 in NO. The best fitting model is reported in Table 6B.

INSERT TABLE 7 ABOUT HERE

To resolve the interaction, we fixed one level of the factor Size to 0 using treatment coding, and refit the model shown in Table 6B twice (once for big, once for small pictures). In addition, we tested for differences between DIFFERENT and NO and differences between SAME and NO separately (i.e., we used treatment coding for the factor Partner, and took NO as the reference level). We did this, instead of using planned contrast coding as in previous analyses, because the interaction was not predicted. When participants named big pictures, they made significantly more errors in the DIFFERENT than in the NO condition (log-odds B = .59, SE=.23, z=2.60), but they made similar amounts of errors in the SAME compared to the NO condition (log-odds B=.15, SE=.25, z=.61). When speakers where naming small pictures, instead, they made marginally fewer errors in the DIFFERENT than in the NO condition (log-odds B = -.37, SE=.21, z=-1.82); again, they made comparable amounts of errors in the SAME as in the NO condition (log-odds B = -.15, SE=.20, z=-.79). So, it appears that in both cases participants' behavior differed when they believed that their partner was naming a different picture compared to when they believed their partner was not naming.

Naming Latencies

Naming latencies longer than 3000 or shorter than 300 ms were considered outliers and removed from all analyses. There were 4 such cases in Experiment 4. Values that were more than 3 standard deviations from the by-participant mean

(1.5%) were replaced with the cut-off value.[8] As shown in Table 7B, by-participant random slopes for Size contributed to model fit ($\chi2$ (5) =12.91, p<.05). Only the main effect of Partner was significant ($\chi2$ (2) =11.20, p<.01). As shown in Table 8 (see also Figure 3, Panel D), latencies were longer when participants believed their partner was responding (naming: 817 ms) than when they did not (NO: 801 ms; Partner 1: B = 15 ms, SE = 5, t = 3.35). However, latencies were no longer in the DIFFERENT (818 ms) than in the SAME condition (817 ms; Partner 2: B = -.01, SE = 4, t = -.13).

INSERT TABLE 8 ABOUT HERE

Of course, this finding contrasts with the results of Experiment 3, where the main effect of Partner was not significant. However, in Experiment 3 we found a non-significant trend in the same direction. To directly compare the two experiments, we ran a combined analysis of Experiment 3 and 4. We found a significant main effect of Partner ($\chi2$ (2) =12.54, p<.005) and no interaction between Partner and Experiment ($\chi2$ (2) =2.15, p=.34). There was (as expected) a main effect of Experiment ($\chi2$ (1) =118.63, p<.001), with latencies being longer when pictures where degraded than when they were not (see Table 8B). [9]

[8] Additional analyses performed on the complete data set (with only the outliers above 3000ms or below 300ms removed) yielded a similar pattern of results.

[9] The combined analysis of Experiments 3 and 4 also revealed a Relatedness*Size interaction ($\chi2$ (1) =4.27, p<.05): a trend towards semantic facilitation when participants named big pictures (B=-5, SE=6, t=-.76), and a reliable semantic interference effect when participants named small pictures (B=13, SE=6, t=2.17; see Table 8B).

Discussion

In Experiment 4, we replicated our finding that naming is slowed down by the belief that one's partner is naming using a task in which participants name single visually degraded pictures. This contrasts with the failure to replicate the effect in Experiment 3, when participants were naming single intact pictures, though we note that the combined analysis did not reveal a difference between the experiments. We therefore cannot be certain whether degradation enhances the interference effect from representing one's partner's task or not.

In Experiment 4, the interaction between Partner and Size in the accuracy analyses was not predicted. The finding that participants produced more errors when naming big pictures in the DIFFERENT than the NO condition (but similar errors in the SAME and the NO conditions) might suggest that they represented the content of their partner's response, as predicted by the task co-representation account. But a similar pattern did not occur with the small pictures, and was not mirrored by the latency analyses, which conformed to the predictions of the actor co-representation account.

General Discussion

Our experiments had pairs of participants naming pictures and manipulated their beliefs about what their partners were doing. We found that naming was slower when participants believed their partners were naming pictures than when they believed their partners were doing nothing or were categorizing pictures. However, their naming was unaffected by whether they believed their partners were performing the same act of naming or a different act of naming. Specifically, when they named two pictures, it did not matter whether they believed their partner was naming those pictures in the same or the reversed order; and when they named a single picture, it did not matter whether their partner named the same or a different picture.

Previous research suggests that participants represent each other's task during non-linguistic joint actions (Knoblich et al., 2011). In Experiments 1, 2, and 4, we found that participants named pictures more slowly when they believed their partners were also naming pictures than when they believed they were doing nothing (Experiments 1 and 4) or were verbally categorizing pictures (Experiment 2). These findings demonstrate that participants represent their partner's task during naming, and therefore indicate that co-representation takes place during language production.

As well as ruling out the no co-representation account, our experiments are incompatible with the task co-representation account. In other words, there was no evidence that participants were representing the content of their partner's response. They were unaffected by whether they believed their partner named the same picture as themselves (Experiments 3 and 4) or a pair of pictures in the same order as themselves (Experiments 1 and 2). In Experiments 1 and 2, beliefs about the content of one's partner response (same vs. different order) also failed to modulate the magnitude of the semantic relatedness effect (the lack of such modulation in Experiments 3 and 4 is harder to interpret given the lack of a main effect of semantic relatedness in those experiments).

Our findings are also incompatible with the actor co-representation account as it is currently formulated (Wenke et al., 2011). Participants were affected by their beliefs about whether their partner was responding or not. However, Experiment 2's results show that interference is specifically due to the belief that one's partner is preparing a *naming* response (as opposed to *any* response). We suggest that interference is (at least partly) due to the belief that another speaker is concurrently engaged in the process of translating a concept into language (i.e., the process of lexical access; Levelt, et al., 1999), and that it is not entirely due to the belief that another speaker is producing a verbal response.

Our experiments therefore showed that speakers represent whether another speaker is concurrently engaged in an act of naming and that doing so interferes with their own act of naming. Therefore, we propose that speakers use their own language production mechanisms to represent whether another speaker is about to produce an utterance. Such a process is a form of internal (covert) simulation of another person's behavior, and is consistent with the claim that beliefs about other people's behavior can drive (or, indeed, are based on) such simulations (e.g., Goldman, 2006). The interference occurs because production and simulated production share some resources.

In one simulation-based account, Pickering and Garrod (2013) proposed that speakers predict their own utterances using "forward models" that are central to theories of motor control and its development (e.g., Wolpert, 1997). Similarly, they predict what their partners are likely to say by covertly imitating their partner's utterances and (in effect) determining what they would say next if they were "in their partner's shoes." But this process can also cause the speaker to engage in some of the processes involved in actual production. When two partners are speaking consecutively, this should not lead to interference, but interference may occur if they speak concurrently. If speakers cannot hear their partners, but believe they are in the process of speaking, they may also predict their partners' utterances and engage production processes. This would lead to interference, even though they do not hear their partners' utterances.

Pickering and Garrod's (2013) account is concerned with the representation of other people's utterances during comprehension. However, the mechanisms can be applied to situations in which the other person's utterance is not heard, as in our experiments. In this case, the participant does not predict on the basis of a partial utterance, but rather on the basis of the belief that his or her partner is preparing to speak at the same time as him (belief which is in turn based on the instructions received by his or her partner). Thus an account developed for the purposes of prediction can be applied to cases of pure imagination (when the partner's utterance is inaudible).

So, why did participants experience interference and take longer to name pictures when they believed their partners were naming pictures themselves, regardless of whether they performed the same or a different act of naming? Neither the task co-representation nor the actor co-representation account can presently accommodate this particular pattern of results. Below, we explain this finding in light of Pickering and Garrod's (2013) suggestion that speakers may covertly simulate others' utterances.

One possibility is that participants formed a representation of their partner's response during covert imitation that specified only the partner's intention to name, but not the content of the partner's upcoming utterance. When the participants ran this representation through their own production systems, at the same time as they were preparing to name (according to their own instructions), interference was caused by a conflict between the speaker's own intention to name and a representation of the partner's intention to name. This could be interpreted as a form of actor conflict, but one that is specific to the task of naming pictures.

Another possibility is that participants ran this representation of their partner's intention to name through their forward production model, thus anticipating the production of an utterance. This forward-model prediction could have triggered the allocation of resources away from the concurrent process of producing the participant's own utterance, thus leading to slower naming latencies when the participants believed their partner was naming at the same time as them. Either version of this account would predict that speakers experience interference whenever they represent that another speaker is about to engage in language production processes (and specifically, in lexical retrieval) at the same time as them.

The finding that speakers represent whether it is their partner's turn to produce language, and that such representation interferes with the production of their own utterances, may help explain the observation that people tend to avoid speaking at the same time (Clark, 1996; Sacks, Schegloff, & Jefferson, 1974; Stivers et al., 2009; see also Schegloff, 2000). Turn-taking is a particularly important component of successful conversations, and knowing when it is one's turn to speak is crucial to avoid overlapping with other speakers. This observation is compatible with our findings that speakers represent when another person is about to name, even in the absence of interaction.

Future research should consider how interference depends on the nature of the stimuli and the nature of the interlocutors. For example, our experiments involved joint picture naming, primarily because we assumed that interference would be most likely when the speaker's believed that both participants were engaged in the same task. But interference occurs in individual picture-word tasks (e.g., Schriefers, Meyer, & Levelt, 1990) and such interference might be affected by beliefs about one's partner. Indeed, some recent evidence suggests that picture-word interference can be eliminated, but only if the speaker believes that her partner is naming a word that is presented in alternating case (Sellaro, et al., 2013). Second, Tsai et al. (2008) showed that joint Simon effects are not present when participants believe their partner is not an intentional agent (i.e., a computer). It may be that interference depends on the speaker believing that another intentional agent is concurrently naming. Third, it is possible that interference is dependent upon the belief that another speaker is producing language at the same time as the participant, and it might be reduced or eliminated if the participant believes that her partner is producing language before or after them.

In conclusion, we have shown that naming responses are affected by the belief that another speaker is concurrently preparing to produce a naming response. This shows that language production mechanisms can be used to represent whether another speaker is about to engage in language production, even in non-interactive contexts.

Acknowledgments

We thank Eddie Dubourg and Ziggy Campbell for assistance with recording equipment, and Sven Radcke for assistance in coding part of the recordings from Experiments 3 and 4. C. Gambi was supported by a University of Edinburgh scholarship. J. Van de Cavey is supported by a FWO scholarship.

References

Almeida, J., Knobel, M., Finkbeiner, M., & Caramazza, A. (2007). The locus of the frequency effect in picture naming: when recognizing is not enough. *Psychonomic Bulletin & Review, 14*(6), 1177-1182.

Aristei, S., Zwitserlood, P., & Abdel Rahman, R. (2012). Picture-induced semantic interference reflects lexical competition during object naming. *Frontiers in Psychology, 3*. doi: 10.3389/fpsyg.2012.00028

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: Sharing tasks with real and imagined co-actors. *Experimental Brain Research, 211*(3-4), 371-385.

Atmaca, S., Sebanz, N., Prinz, W., & Knoblich, G. (2008). Action co-representation: The joint SNARC effect. *Social Neuroscience, 3*(3), 410-420.

Audacity. (Version 1.2.5) [Computer Program]: http://audacity.sourceforge.net.

Audition. (Version 4.0) [Computer Software]. San Jose, CA: Adobe Systems Incorporated. Retrieved from www.adobe.com

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278.

Bates, D., Maechler, M., & Dai, B. (2008). Lme4: Linear mixed-effects models using S4 classes. (Version 0.999375-28).

Bayeen, R. H., Davidson, D. J., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Clark, H. H. (1996). *Using language*. Cambridge, U.K.: Cambridge University Press.

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review, 10*(1), 111-117.

E-Prime. (Version 2.0) [Computer Software]. Pittsburgh, PA: Psychology Software Tools. Retrieved from www.pstnet.com

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 640-654.

Goldman, A. (2006). *Simulating minds*. New York: Oxford University Press.

Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: Estimating Lognormal, Wald, and Weibull RT distributions with a parameter-dependent lower bound. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 277-290.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review*, *9*(2), 394-401.

Heyes, C. (2011). Automatic imitation. *Psychonomic Bulletin & Review, 137*(3), 463-483.

Hommel, B., Colzato, L. S., & van den Wildenberg, W. P. M. (2009). How Social Are Task Representations? *Psychological Science, 20*(7), 794-798.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*(05), 849-937.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434-446.

Jaeger, T. F., Furth, K., & Hilliard, C. (2012). Incremental phonological encoding during unscripted sentence production. *Frontiers in Psychology*, *3*:481, doi: 10.3389/fpsyg.2012.00481

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 824-843.

Klauer, K. C., Herfordt, J., & Voss, A. (2008). Social presence effects on the Stroop task: Boundary conditions and an alternative account. *Journal of Experimental Social Psychology*, *44*(2), 469-476.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and Data. In B. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 54, pp. 59-101). Burlington: Academic Press.

La Heij, W., Heikoop, K. W., Akerboom, S., & Bloem, I. (2003). Picture naming in picture context: Semantic interference or semantic facilitation? *Psychology Science, 45*(1), 49-62.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-75.

Mädebach, A., Jescheniak, J. D., Oppermann, F., & Schriefers, H. (2011). Ease of processing constraints the activation flow in the conceptual-lexical system during speech planning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(3), 649-660.

Meyer, A. S., & Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory & Cognition, 35*(3), 494-503.

Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 555-563.

Navarrete, E., & Costa, A. (2005). Phonological activation of ignored pictures: Further evidence for a cascade model of lexical access. *Journal of Memory and Language, 53*(3), 359-377.

Philipp, A. M., & Prinz, W. (2010). Evidence for a role of the responding agent in the joint compatibility effect. *Quarterly Journal of Experimental Psychology, 63*(11), 2159-2171.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(2), 169-226.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329-392. doi: http://dx.doi.org/10.1017/S0140525X12001495

Pickering, M. J., & McLean, J. F. (2013). *Representing others' words: Just like one's own?* Unpublished manuscript. University of Edinburgh. Edinburgh, U.K.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*(2), 129-154.

R. (Version 2.13.1) [Computer Program]. Vienna, Austria: R Development Core Team, http://www.R-project.org.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510.

Ruys, K. I., & Aarts, H. (2010). When competition merges people's behavior: Interdependency activates shared action representations. *Journal of Experimental Social Psychology, 46*(6), 1130-1133.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*(4), 696-735.

Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society, 29*(1), 1-63.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language, 29*(1), 86-102.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences 10*(2), 70-76.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88*(3), B11-B21.

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1234-1246.

Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006). Twin peaks: An ERP study of action planning and control in coacting individuals. *Journal of Cognitive Neuroscience, 18*(5), 859-870.

Sellaro, R., Treccani, B., & Cubelli, R. (2013). *When task sharing eliminates interference: Evidence from the joint Picture-Word interference paradigm*. Poster presented at the 5th Joint Action Meeting, Berlin.

Smith, M., & Wheeldon, L. (2004). Horizontal information flow in spoken language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 675-686.

Snodgrass, J. G., & VanderWart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 174-215.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1280.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., . . . Levinson, S. C. (2009). Universals and cultural variation in turn--taking in conversation. *Proceedings of the National Academy of Sciences, 106*(26), 10587–10592.

Tsai, C.-C., Kuo, W.-J., Jing, J.-T., Hung, D. L., & Tzeng, O. J. L. (2006). A common coding framework in self-other interaction: Evidence from joint action task. *Experimental Brain Research, 175*(2), 353-362.

Tsai, C.-C., Kup, W.-J., Hung, D. L., & Tzeng, O. J. L. (2008). Action co-representation is tuned to other humans. *Journal of Cognitive Neuroscience, 20*(11), 2015-2024.

Van de Cavey, J., Gambi, C., MacKenzie, A., Nelissen, K., & Pickering, M. J. (2012, September). *What do speakers represent about others' utterances? Comparing simultaneous and asynchronous production*. Poster presented at the AMLaP (Archietectures and Mechanisms for Language Processing) conference. University of Trento. Riva del Garda, Italy.

Vlainic, E., Liepelt, R., Colzato, L. S., Prinz, W., & Hommel, B. (2010). The virtual co-actor: The social Simon effect does not rely on online feedback from the other. *Frontiers in Psychology, 1*, doi: 10.3389/fpsyg.2010.00208.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(2), 423-440.

Wenke, D., Atmaca, S., Holländer, A., Liepelt, R., Baess, P., & Prinz, W. (2011). What is shared in joint action? Issues of co-representation, response conflict, and agent identification. *Review of Philosophy and Psychology, 2*(2), 147-172.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin, 131*(3), 460-473.

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences, 1*(6), 209-216.

Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 255-260.

Appendix A

Experimental materials used in Experiments 1-4

INSERT TABLE 1A ABOUT HERE

Appendix B

Linear mixed-effects results for Experiments 1-4, and the combined analyses of Experiments 3 and 4.

Experiment 1.

Accuracy. The best-fitting model corresponding to analyses reported in text is shown in Table 1B. Additional maximal random structure LME analyses (with the correlation between random intercepts and random slopes fixed to zero) further confirmed these results. In the complete model (with all fixed effects parameters), the only coefficients significantly different from zero were the ones associated with the two planned contrasts for Partner (Partener1: B = .33, SE = .14, z=2.36; Partner2: B = -.27, SE=.09, z=-3.13; all other p's <.1)[10].

INSERT TABLE 1B ABOUT HERE

Naming Latencies. The best-fitting model corresponding to analyses reported in text is shown in Table 2B. LME analyses with maximal-random structure (with the correlation between random intercepts and random slopes fixed to zero, and the 3-way interaction of Size, Partner, and Relatedness by-participants removed to aid convergence) further confirmed these results. In the complete model (with all fixed effects parameters), the only coefficients significantly different from zero were the one associated with the first planned contrasts for Partner (Partener1: B = 13 ms, SE = 5, t=2.77; Partner2: B = 1 ms, SE=5, t=0.16) and the one associated with the factor Relatedness (B: 16 ms, SE = 6, t= 2.73; all other |t|'s <1.9)[11].

[10] Coefficients were taken to be significantly different from zero if |z|>2.
[11] Coefficients were taken to be significantly different from zero if |t|>2.

Additional LME analyses were carried out on inverse transformed naming latencies. Such additional analyses used the maximal random structure that converged (see above). In the complete model (with all fixed effects parameters), the coefficient associated with the first planned contrasts for Partner (Partener1: B = -19, SE = 6, t=-3.33; Partner2: B = 3, SE=5, t=0.50) and the one associated with the factor Relatedness (B: -20, SE = 7, t= -2.83) were both significant.

INSERT TABLE 2B ABOUT HERE

Experiment 2.

Accuracy. LME analyses with full-random structure could not be carried out due to convergence issues. However, a model with by-participant and by-items random slopes for both Relatedness and Partner (but not their interaction) confirmed that no factor contributed to model fit. In the complete model (with all fixed effects parameters), no coefficient was significantly different from zero (all |z|'s <1.6).

Naming latencies. The best-fitting model corresponding to analyses reported in text is shown in Table 3B. LME analyses with full-random structure could not be carried out due to convergence issues. The most complex random structure that converged contained random slopes for the factors of interest (Partner, Relatedness, and their interaction), but not for the factor Size (nor for any interactions involving Size; we also fixed the correlation between random intercepts and random slopes to zero, as in previous analyses). In the complete model (with all fixed effects parameters), the only coefficients significantly different from zero were the one associated with the first planned contrasts for Partner (Partener1: B = 12 ms, SE = 6, t=2.16; Partner2: B = 3 ms, SE=6, t=0.50) and the one associated with the factor Relatedness (B: 19 ms, SE = 6, t=3.36; all other |t|'s <1.6).

Additional LME analyses were carried out on inverse-transformed naming latencies. Such additional analyses used the maximal random structure that converged. In the complete model (with all fixed effects parameters), the coefficient associated with the first planned contrasts for Partner was not reliable (Partener1: B = -12, SE = 7, t=-1.75; Partner2: B =-1, SE=7, t=-0.22), while the one associated with the factor Relatedness was significant (B: -23, SE = 7, t= -3.36).

In order to better understand the reasons for this discrepancy between analyses with and without the inverse transformation, we conducted ex-Gaussian analyses. The inverse transformation is often used to normalize the distribution of the data. In our case, the residuals of LME models applied to inverse-transformed data are indeed closer to a normal distribution than the residuals of models applied to untransformed data. Ratcliff (1993) showed that the inverse transformation normally preserves power well. However, his simulations using the ex-Gaussian distribution showed that there is one case in which the inverse transformation is associated with a loss in the power to detect real effects (see Figure 5, p. 517). This is when the effect is in the τ parameter rather than in the μ parameter of the ex-Gaussian, and across-subjects variability is large (300 ms range in by-participant means in his simulations).

We fitted an ex-Gaussian function to the raw data from each participant in each of the six cells in our design, obtained by fully crossing the factors of interest, Partner and Relatedness. Size was ignored to maximize the performance of the fitting algorithm, which is known to compute reliable estimates when the number of observations per subject per cell is at least 40 (Heathcote, Brown, & Mewhort, 2002). In our analyses the number of observations ranged from 39 to 50. We used the maximum number of quantiles method provided in QMPE (Heathcote, Brown, & Cosineau, 2004), which has been shown to provide the best fits (Heathcote et al., 2002; see also Staub, White, Drieghe, & Hollway, 2010 for an application to psycholinguistic research). We obtained estimates for the τ and μ parameters, and entered them into by-participant analyses of variance with Partner and Relatedness as within-participants factors. We also estimated across-participants variability by computing the range of by-participants means.

For Experiment 2, we found no effect on the μ parameter (all F's < 1.4, all p's >.200). Instead, we found a main effect of Relatedness on the τ parameter ($F(1;31) = 7.275$, $p<.05$). Crucially, we also found a main effect of Partner on the τ parameter ($F(2;62) = 3.306$, $p<.05$). The interaction was not significant ($p= .174$). See Table 4B. The range in by-participant means was 581 ms in Experiment 2. This is not only large compared to Ratcliff's (1993) simulations, but also larger than in Experiment 1 (where the range is 368 ms). Therefore, we conclude that the inverse transformation should not be applied to data from Experiment 2. Furthermore, ex-Gaussian analyses confirmed that Partner has an effect on the distribution of naming latencies in this Experiment and, crucially, such analyses do not assume normality (to the contrary, they explicitly model non-normality).

INSERT TABLE 3B ABOUT HERE

INSERT TABLE 4B ABOUT HERE

Experiment 3.

Accuracy. LME analyses with maximal random structure could not be carried out due to convergence issues. However, a model with by-participant and by-items random slopes for both Relatedness and Partner (but not their interaction) showed that no coefficient differed significantly from zero (all $|z|$'s <2).

Naming Latencies. The best-fitting model corresponding to analyses reported in text is shown in Table 5B. LME analyses with full-random structure could not be carried out due to convergence issues. The most complex random structure that converged contained by-participant random slopes for Size, Partner, and Relatedness (but no interaction), and by-items random slopes for Partner and Relatedness (but not their interaction; we also fixed the correlation between random intercepts and random slopes to zero, as in previous analyses). In the complete model (with all fixed effects parameters), only the Relatedness-by-Size interaction reached significance ($B = 22$, $SE = 10$, $t = 2.18$). Simple main effect analyses confirmed that when participants were naming big pictures, naming latencies showed a non-significant tendency towards semantic facilitation ($B = -8$, $SE = 7$, $t = -1.14$); when participants were naming small pictures, instead, they showed a tendency towards semantic interference ($B = 14$, $SE = 7$, $t = 1.94$).

Additional LME analyses were carried out on inverse-transformed naming latencies. Such additional analyses used the largest random structure that converged (see above). As in the analyses with untransformed latencies, in the complete model (with all fixed effects parameters) only the coefficient for the Relatedness-by-Size interaction was significant ($B = -46$, $SE = 21$, $t = -2.13$). Simple main effects analyses confirmed the non-significant trends described in the previous paragraph (big pictures, Relatedness: B = 20, SE = 15, t = 1.34; small pictures, Relatedness: B = -25, SE = 15, t = -1.68).

INSERT TABLE 5B ABOUT HERE

Experiment 4.

Accuracy. The best-fitting model corresponding to analyses reported in text is shown in Table 6B. LME analyses with full-random structure could not be conducted due to convergence issues. The most complex random structure that converged contained by-participant random slopes for Partner, Relatedness, Size, the Partner-by-Relatedness interaction, and the Partner-by-Size interaction, and by-items random slopes for Partner, Relatedness, and the Partner-by-Relatedness interactions (all random correlations were fixed to zero to aid convergence). The complete model (with all fixed effects parameters) showed that only the interaction of Partner and Size was significant (Partner1*Size, B = -.91, SE =.29, z = -3.13; Partner2*Size, B = .68, SE = .24, z = 2.91). Simple main effects analyses could not be conducted due to convergence issues.

INSERT TABLE 6B ABOUT HERE

Naming latencies. The best-fitting model corresponding to analyses reported in text is shown in Table 7B. LME analyses with full-random structure could not be carried out due to convergence issues. The most complex random structure that converged contained by-participant random slopes for Size, Partner, and Relatedness (but no interaction), and by-items random slopes for Partner and Relatedness (but not their interaction; we also fixed the correlation between random intercepts and random slopes to zero). In the complete model (with all fixed effects parameters), only the coefficient associated with Partner1 was significant (Partner1, B = 16, SE = 5, t = 3.03), thus confirming that participants took longer to name when they believed their partner was also naming pictures than when they believed their partner was silent. Instead, they took no longer to name pictures when they believed their partner was naming a different picture than when they believed their partner was naming the same picture (Partner2: B = -1, Se = 4, t = -.12).

Additional LME analyses were carried out on inverse-transformed naming latencies. Such additional analyses used the largest random structure that converged (see above). As in the analyses with untransformed latencies, in the complete model (with all fixed effects parameters) only the coefficient associated with Partner1 was significant (Partner1, B = -19, SE = 7, t = -2.82), but not the coefficient associated with Partner 2 (Partner 2, B = 2, Se = 6, t = .30). No other main effect or interaction was significant.

INSERT TABLE 7B ABOUT HERE

Combined analyses of Experiment 3 and Experiment 4.

Naming latencies. The best-fitting model corresponding to analyses reported in text is shown in Table 8B. LME analyses with full-random structure could not be carried out due to convergence issues. The most complex random structure that converged contained by-participant random slopes for Size, Partner, and Relatedness (but no interaction), and by-items random slopes for Experiment, Partner and Relatedness (but not their interaction; we also fixed the correlation between random intercepts and random slopes to zero, as in previous analyses). In the complete model (with all fixed effects parameters), the coefficient associated with Partner1 was significant (Partner1, $B = 11$, $SE = 4$, $t = 2.68$), whereas the coefficient associated with Partner 2 was not (Partner2, $B = 1$, $SE = 3$, $t = .36$). In addition, there was a main effect of Experiment ($B = 112$, $SE = 5$, $t = 21.68$) and also a Relatedness-by-Size interaction ($B = 18$, $SE = 9$, $t = 2.08$). Simple main effect analyses showed a non-significant trend towards semantic facilitation for big pictures ($B = -5$, $SE = 6$, $t = -.76$) and a significant interference effect for small pictures ($B = 13$, $SE = 6$, $t = 2.16$).

LME analyses on inverse-transformed naming latencies (with the same random structure as the analyses mentioned above) further confirmed these results: Partner 1 ($B = -16$, $SE = 7$, $t = -2.39$), Partner 2 ($B = -1$. $SE = 5$, $t = -.16$), Experiment ($B = -199$, $SE = 8$, $t = -24$), Relatedness-by-Size ($B = -31$, $SE = 15$, $t = -2.00$). Simple main effects analyses cannot be reported due to convergence issues.

INSERT TABLE 8B ABOUT HERE

Table 1.

Error counts (percentages) by Partner and Relatedness in Experiment 1.

	DIFFERENT	SAME	NO
Unrelated	95 (7.9%)	81 (6.8%)	76 (6.3%)
Related	97 (8.1%)	64 (5.3%)	59 (4.9%)

Table 2.

Mean voice onset times in ms (and standard deviations) by Partner and Relatedness in Experiment 1.

	DIFFERENT	SAME	NO	Tot
Unrelated	869 (220)	869 (225)	855 (217)	864
Related	881 (221)	886 (230)	872 (229)	880
Tot	875	877	864	

Table 3.

Error counts (percentages) by Partner and Relatedness in Experiment 2.

	DIFFERENT	SAME	CAT
Unrelated	89 (5.6%)	100 (6.3%)	96 (6.0%)
Related	115 (7.2%)	114 (7.1%)	93 (5.8%)

Table 4.

Mean voice onset times in ms (and standard deviations) by Partner and Relatedness in Experiment 2.

	DIFFERENT	SAME	CAT	Tot
Unrelated	881 (257)	879 (257)	874 (255)	878
Related	898 (257)	907 (277)	885 (259)	897
Tot	889	893	880	

Table 5.

Error counts (percentages, out of the total number of scorable trials) by Partner and Relatedness in Experiment 3.

	DIFFERENT	SAME	NO
Unrelated	45 (3.8%)	38 (3.2%)	31 (2.6%)
Related	31 (2.6%)	46 (3.8%)	37 (3.1%)

Table 6.

Mean voice onset times in ms (and standard deviations) by Partner, Relatedness, and Size in Experiment 3.

	DIFFERENT	SAME	NO	
		Unrelated		
Big	694 (174)	698 (171)	691 (168)	
Small	709 (148)	710 (176)	697 (152)	
Tot unrelated	701	704	694	700
		Related		
Big	688 (175)	692 (172)	681 (166)	
Small	716 (164)	722 (168)	719 (169)	
Tot related	702	707	700	703
	DIFFERENT	SAME	NO	
Tot	702	705	697	

Table 7.

Error counts (percentages, out of the total number of scorable trials) by Partner, Relatedness, and Size in Experiment 4.

	DIFFERENT	SAME	NO	
		Unrelated		
Big	33 (5.5%)	21 (3.5%)	15 (2.5%)	
Small	24 (4.0%)	28 (4.7%)	32 (5.3%)	
Tot unrelated	57	49	47	153
		Related		
Big	25 (4.2%)	18 (3.0%)	19 (3.2%)	
Small	24 (4.0%)	30 (5.0%)	34 (5.7%)	
Tot related	49	48	53	150
	DIFFERENT	SAME	NO	
Tot	106	97	100	

Table 8.

Mean voice onset times in ms (and standard deviations) by Partner and Relatedness in Experiment 4.

	DIFFERENT	SAME	NO	Tot
Unrelated	817 (208)	813 (215)	798 (198)	809
Related	818 (222)	820 (214)	805 (206)	815
Tot	818	817	801	

Table 1A.

Big pictures, their semantic categories, and the small pictures they were paired with in the unrelated and related conditions.

Big	Small - unrelated	Small-related	Semantic category
apple	blouse	banana	food
bed	dress	chair	furniture
boat	leg	plane	means of transport
bowl	cake	vase	container
bread	guitar	cake	food
cap	vase	dress	clothing
car	seal	helicopter	means of transport
cat	pan	fish	animal
drum	table	guitar	music instrument
foot	pig	ear	body part
frog	banana	seahorse	animal
glass	waistcoat	cup	container
goat	cup	pig	animal
hand	seahorse	eye	body part
horse	trousers	seal	animal
jug	chair	bottle	container
knife	helicopter	pan	utensil
nose	plane	leg	body part
onion	ear	carrot	food
pear	bottle	grapes	food

saw	eye	pliers	utensil
shoe	fish	trousers	clothing
skirt	grapes	waistcoat	clothing
sock	carrot	blouse	clothing
sofa	pliers	table	furniture

Table 1B.

Best fitting model for the accuracy data in Experiment 1.

Predictor	Estimate	SE	Z
Intercept	-3.10	.18	-16.97
Partner1: naming vs. NO	.24	.11	2.23
Partner2: SAME vs. DIFFERENT	-.23	.08	-2.75
Random effect	**Variance estimate**		
Subjects: intercept	.48		
Items: intercept	.48		
Items: Relatedness	.56		

Table 2B.

Best fitting model for the voice onset time data in Experiment 1.

Predictor	Estimate	SE	t
Intercept	874	24	36.72
Partner1: naming vs. NO	14	5	2.79
Partner2: SAME vs. DIFFERENT	1	4	.17
Relatedness (related vs. unrelated)	16	5	3.36
Random effect	Variance estimate		
Subjects: intercept	11980		
Items: intercept	3150		

Note. Variance estimates have been rounded up to the nearest ten in all the analyses of naming latencies reported in this paper (the analyses were run on latencies measured in seconds, and then the estimates were converted to milliseconds for expository purposes).

Table 3B.

Best fitting model for the voice onset time data in Experiment 2.

Predictor	Estimate	SE	t
Intercept	884	24	36.77
Partner1: naming vs. CAT	12	5	2.47
Partner2: SAME vs. DIFFERENT	3	4	.70
Relatedness (related vs. unrelated)	19	5	3.48
Random effect	Variance estimate		
Subjects: intercept	16490		
Subjects: Size	13080		
Items: intercept	46670		
Items: Relatedness	4380		

Table 4B.

Mean values of the τ parameter in ms by Partner and Relatedness in Experiment 2.

	DIFFERENT	SAME	CAT	Tot
Unrelated	220	208	204	211
Related	224	241	213	226
Tot	222	225	209	

Table 5B.

Best fitting model for the voice onset time data in Experiment 3.

Predictor	Estimate	SE	t
Intercept	702	18	39.87
Partner1: naming vs. NO	7	4	1.88
Partner2: SAME vs. DIFFERENT	3	3	.91
Relatedness (related vs. unrelated)	3	5	.59
Size (small vs. big)	21	16	1.34
Relatedness: Size	22	10	2.19
Random effect	**Variance estimate**		
Subject: intercept	5990		
Subject: Size	320		
Item: intercept	2880		
Item: Relatedness	760		

Table 6B.

Best fitting model for the accuracy data in Experiment 4.

Predictor	Estimate	SE	Z
Intercept	-3.88	.21	-18.18
Partner1: naming vs. NO	.05	.13	.39
Partner2: SAME vs. DIFFERENT	-.08	.11	-.73
Size: small vs. big	.10	.40	.24
Partner1: Size	-.63	.27	-2.35
Partner2: Size	.47	.21	2.18
Random effect	Variance estimate		
Subject: intercept	.39		
Subject: Size	.98		
Item: intercept	1.10		

Table 7B.

Best fitting model for the voice onset time data in Experiment 4.

Predictor	Estimate	SE	t
Intercept	816	21	39.56
Partner1: naming vs. NO	15	5	3.35
Partner2: SAME vs. DIFFERENT	-.01	4	-.13
Random effect	**Variance estimate**		
Subject: intercept	8090		
Subject: Size	690		
Item: intercept	4220		

Table 8B.

Best fitting model to latency data in Experiments 3 and 4 (combined analysis).

Predictor	Estimate	SE	t
Intercept	759	13	57.56
Partner1: naming vs. NO	11	3	3.51
Partner2: SAME vs. DIFFERENT	1	3	.47
Relatedness: related vs. unrelated	4	4	1.00
Size: small vs. big	22	16	1.39
Experiment: 4 vs. 3	112	5	21.74
Relatedness: Size	18	9	2.07
Random effect	**Variance estimate**		
Subject: intercept	2140		
Subject: Size	1100		
Item: intercept	2790		
Item: Relatedness	530		
Item: Experiment	890		

Note. By-participant random slopes for Size ($\chi2$ (3) =13.91, p<.005); by-items random slopes for Relatedness ($\chi2$ (3) =19.57, p<.001) and Experiment ($\chi2$ (3) =45.42, p<.001). These slopes were selected using a forward method since very complex models would not converge due to the number of factors involved. The model did not include correlations among random slopes and intercepts, unlike previous analyses. For Experiment, the following coding scheme was used: Experiment 4: 1/2, Experiment 3: -1/2.

Figure 1. Partner's task manipulation in Experiment 1 (sample trial from the unrelated condition; *apple* is displayed in blue, *blouse* is displayed in red).

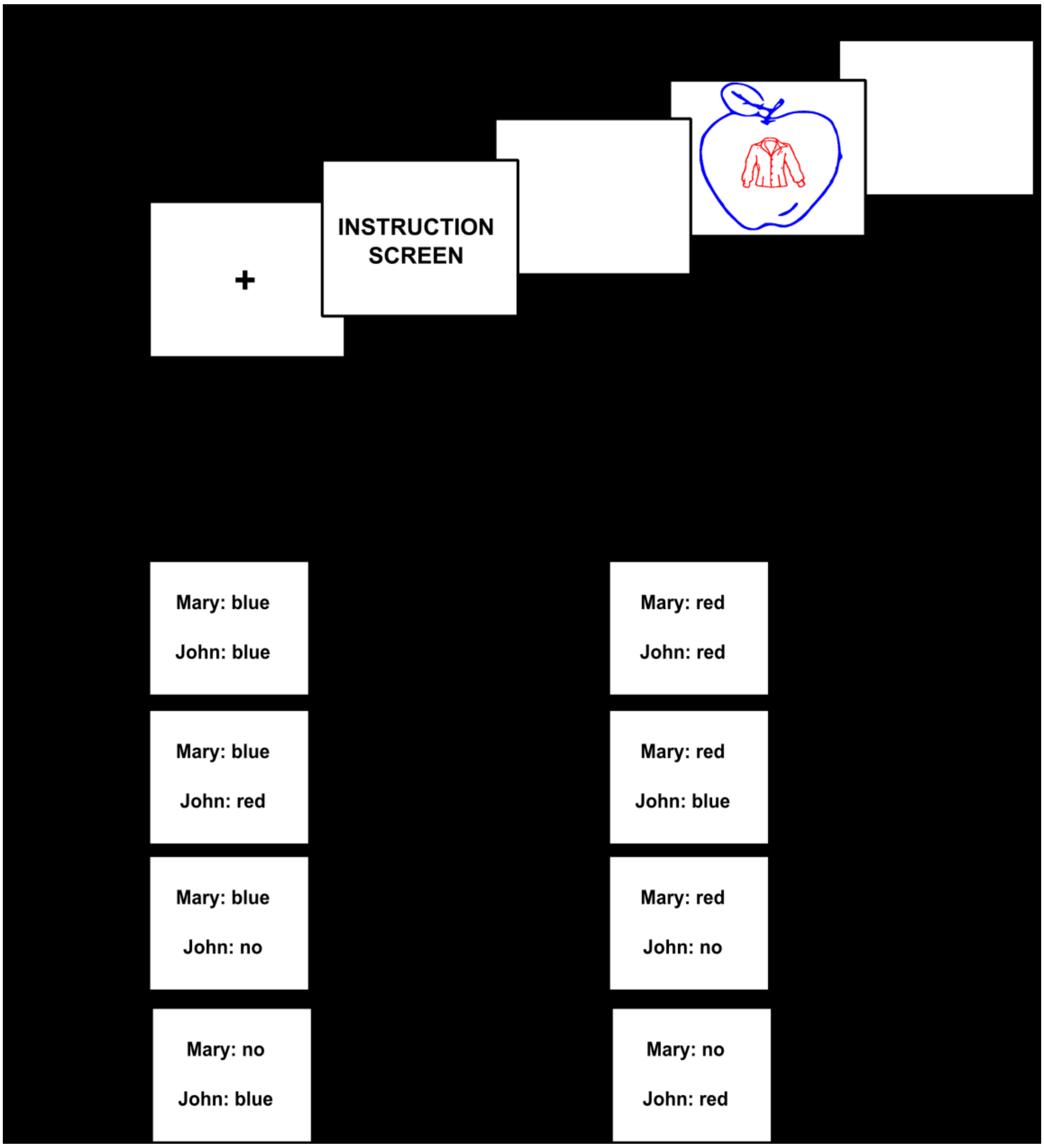

Figure 2. Hypothesized effects of Partner's task according to three different accounts: (A) task co-representation account; (B) actor co-representation account; (C) no co-representation account; *apple* is in blue, *blouse* is in red. The left-hand side of the figure presents static depictions of the relevant nodes in Mary's mental lexicon, with nodes making up Mary's self-representations at the top and nodes making up Mary's representation of John's response (that is, *Mary(John)*) at the bottom. The right-hand side of the figure presents a snapshot of the activation level of the nodes just before the onset of the word "apple" when Mary is preparing to utter "apple blouse" (unrelated condition) and John is either preparing to utter "blouse apple" (DIFFERENT), "apple blouse" (SAME), or nothing (NO). The degree of activation is indicated by the thickness of the circles. Pointed arrows indicate facilitation, while rounded arrows indicate competition. In the related case, *banana* would replace *blouse* and the strength of competition between *banana* and *apple* would be greater than in the depicted unrelated case. Bar graphs indicate the expected pattern of results under each of the accounts.

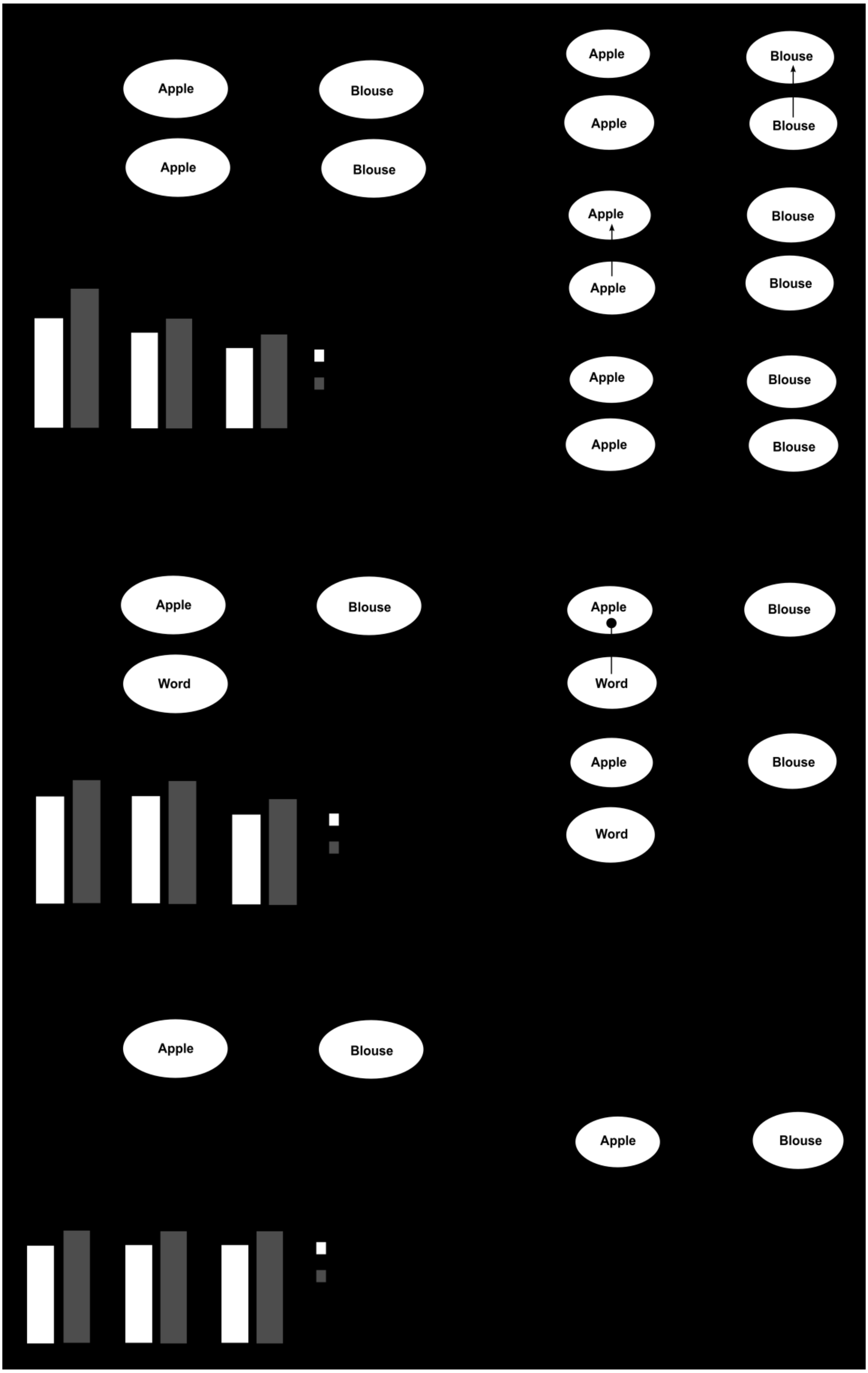
Apple
Blouse
Apple
Blouse
Apple
Blouse
Apple
Blouse
Apple
Blouse
Apple
Blouse
Apple
Blouse
Apple
Blouse
Apple
Blouse
Word
Apple
Blouse
Word
Apple
Blouse
Word
Apple
Blouse
Apple
Blouse

Figure 3. Mean naming onset latencies in ms for Experiments 1-4, by Partner and Relatedness. Bars represent standard errors of the mean.